Redpaper

Vinay Kulkarni

The MAX5 Advantage: How IBM System x MAX5 Benefits Microsoft SQL Server Data Warehouse Workloads

This IBM® Redpaper™ publication describes the series of tests performed by the IBM System x® Performance Lab in Kirkland, Washington, that demonstrate the benefits of using IBM System x3850 X5 with MAX5 for a large-scale decision support system (DSS) workload.

This paper includes the following topics:

Overview

IBM developed the System x MAX5 offering after conducting a research study on customer pain points and learning that insufficient memory was often the source of performance bottlenecks, before processors or the input/output (I/O) subsystem became the bottleneck.

In the previous generation of servers, a single 4-socket server only scaled up to 256 GB of memory. If you ran out of memory at 256 GB, there was no way to just upgrade the memory. You had to attach another 4-socket server that gave you additional unnecessary processors and other components, and it was expensive to do so. With IBM eX5 technology, you have four times the memory capacity (1 TB) on a 4-socket server, plus the ability to add another 512 GB with one rack unit (1U) memory drawer called the MAX5, at a relatively lower cost.

The IBM MAX5 memory drawer is a great option to reduce the I/O footprint of a data warehouse configuration in your data center. Additional savings for space, power, and cooling, also apply, making it even more attractive in terms of price and performance.

For this series of tests, a 1500 GB data warehouse under the control of Microsoft® SQL Server 2008 R2 was used on an IBM System x3850 X5 server configured with four 8-core processors and IBM MAX5 for System x. Tests were conducted with and without MAX5 to measure the performance benefit of MAX5 and additional memory. For storage, IBM High IOPS SSD PCIe adapters were used.

The IBM MAX5 for System x increased the queries per hour of the x3850 X5 by 28%, and decreased average query response times by 39%.

Understanding the DSS workload

A data warehouse is a type of database that is designed for archiving data for the purposes of reporting and analysis. Data warehouses are typically used as the back-end data storage for DSS.

A data warehouse environment is unlike an online transaction processing (OLTP) database in that it is often subjected to ad hoc queries, rather than the predefined queries that an OLTP database is designed to accommodate. Queries that are run against data warehouse databases often scan millions of rows, as opposed to typical OLTP queries, which generally scan a handful of records. The I/O pattern is mostly large sequential reads compared to small random reads with an OLTP workload.

IBM System x3850 X5

The x3850 X5 is built on the fifth generation of IBM X-Architecture® chip set technology. This new system incorporates the latest in hardware acceleration and scalability. These advances provide industry-leading flexibility, with resource configurations that push the boundaries of x86 and x64-based systems.

Figure 1 shows the x3850 X5.

Figure 1 Front view of the x3850 X5 showing eight 2.5-inch SAS drives

With the release of new multicore processors from Intel®, the 4-socket server IBM System x3850 X5 offers up to 32 cores or 64 logical processors with Hyper-Threading Technology and up to 1 TB of memory per chassis. The memory is highly available with Memory ProteXion, Chipkill, and memory mirroring.

Other high-availability features, such as redundant fans, hot-swap fully redundant power supplies with 220V AC input, and hot-swap disk drives, deliver high levels of reliability and availability to maximize up-time. These features provide the reliability and high availability needed for the most demanding database workloads.

The x3850 X5 features four times the memory availability than the previous generation with 64 dual inline memory module (DIMM) slots, running PC3-10600 DDR3 memory.

This paper provides information about how you can use the memory capabilities of the x3850 X5 to help take the performance, efficiency, and reliability of your data warehouse environment to the next level.

IBM MAX5 for System x

The x3850 X5 server can be expanded with the MAX5 1U memory expansion unit to scale the memory up to 96 DIMMs. Using 16 GB DIMMs, a 4-processor x3850 X5 with MAX5 can support a total of 1.5 TB memory.

Figure 2 shows the x3850 X5 and MAX5 together.

Figure 2 - IBM System x3850 X5 with MAX5

MAX5 is an industry-first technology from IBM that separates memory from the processor, eliminating the need to buy another server to support memory-intensive workloads and dramatically changes the economics of the x86 server market. The MAX5 connects to the x3850 X5 with four QuickPath Interconnect (QPI) links, one from each of the four sockets.

Figure 3 shows the logical connections between the MAX5 and x3850 X5.

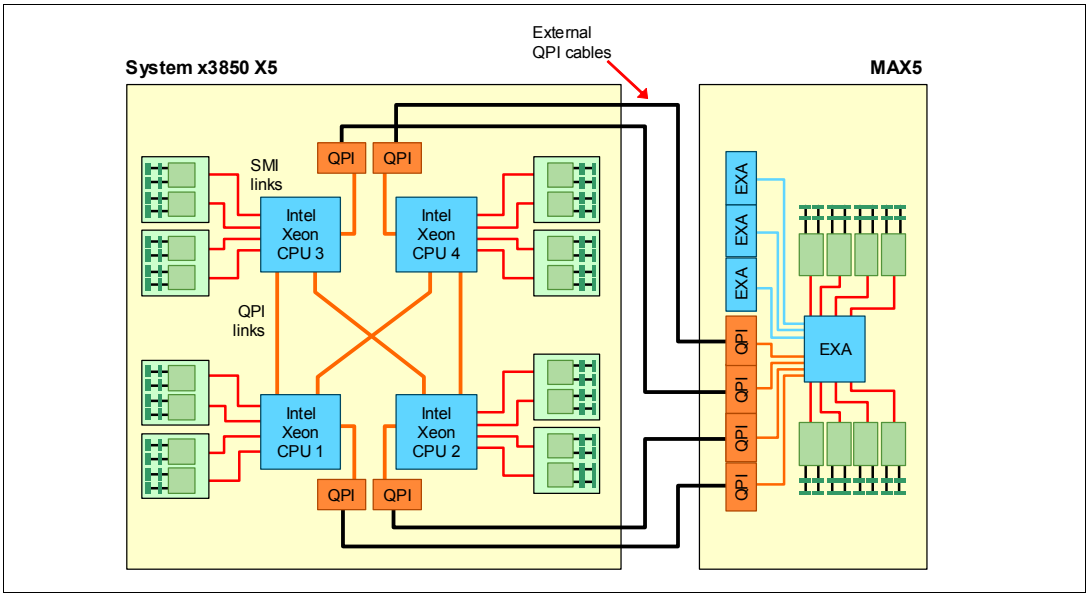


Figure 3 The x3850 X5: connectivity of the system unit with the MAX5

MAX5 is not suitable in all scenarios. If you have less than 256 GB memory in your x3850 X5 and need to upgrade memory, then buying larger size DIMMs can be a more economical and better performing option. If you have maxed out the memory on x3850 X5 at 1 TB and you still have an I/O bottleneck that is slowing down your data warehouse queries, the MAX5 can be an ideal fit for your configuration. Adding a MAX5 into that configuration can be a better option than trying to add more storage into the configuration. The MAX5 can help provide lower latencies than even solid state device (SSD) storage. If you end up caching your database entirely into memory, you can see the additional performance gain.

Microsoft SQL Server 2008 R2

With Microsoft SQL Server 2008 R2 running on Microsoft Windows® Server 2008 R2, you can build comprehensive, enterprise-scale analytic solutions that deliver intelligence where you want it.

Microsoft's SQL Server 2008 R2 makes a major advance in data warehouse scalability. The database engine contains numerous enhancements, such as increased parallelism, improvements in compression, star join query optimizations, and grouping sets for queries. These enhancements were all designed to improve both absolute performance and multi-CPU scaling on decision support workloads.

IBM High IOPS SSD PCIe adapters

The IBM High IOPS SSD PCIe adapters provide a new generation of ultra-high-performance storage for System x servers, based on SSD technology. These adapters provide alternatives to traditional serial-attached SCSI (SAS) disk drives by being able to deliver throughput of up to 100,000 I/O operations per second (IOPS). They are designed for high-performance servers and computing appliances, and help to provide the added benefits of lower power, cooling, management overhead, and a smaller storage footprint.

Figure 4 shows the IBM High IOPS SSD PCIe adapter.

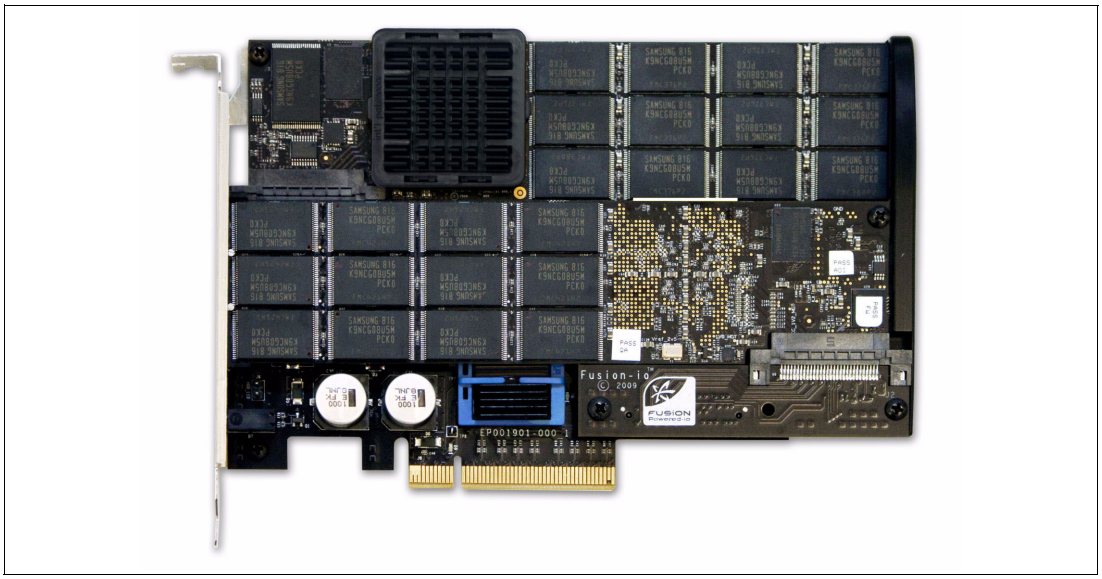


Figure 4 IBM High IOPS SSD PCIe adapter

High IOPS SSD PCIe adapters can help provide the following benefits:

- Doubles the performance of the industry-leading first generation ioDrive.
- Integrates with servers at the system bus and kernel levels, creating a new flash memory tier.
- Outperforms dozens of SSDs in a single server.
- Delivers I/O performance of thousands of SAS hard disk drives in a single server, up to 100,000 IOPS for the 160 GB and 320 GB adapters, and 200,000 IOPS at 1.5 GB for the 640 GB adapter.
- Accelerates applications, improves response times, and boosts efficiency.
- Reduces storage latencies and eliminates I/O bottlenecks.
- Provides from 160 GB to 1.2 TB of enterprise-grade, solid-state flash storage.
- Delivers extremely high reliability, due to advanced wear-leveling, 11-bit error correction code (ECC), and N+1 chip-level redundancy.

You can find more information in the *IBM High IOPS SSD PCIe Adapters at-a-glance guide*, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0729.html?Open

Test environment

Decision support workloads are frequently I/O intensive. The high I/O requirements of these tests were met using High IOPS SSD PCIe adapters installed in the PCI Express (PCIe) slots on a x3850 X5 server.

The test environment consisted of a single x3850 X5 server with four Intel Xeon® X7560 8-core processors. The server was populated with the following High IOPS SSD PCIe adapters:

- Five 320 GB High IOPS SD Class PCIe adapters
- Two 640 GB High IOPS MLC Duo adapters

With these SSD PCIe adapters, the server did not need any external storage to house the 1.5 TB database, making it a complete single-box environment. These adapters offer up to 1.5 GBs large sequential read bandwidth per card and can be ideal for the large sequential read bandwidth demands of a data warehouse workload. They can also be an ideal option for the low latency requirements of a SQL Server tempdb database for this workload.

The following memory configurations were evaluated:

- 1 TB of memory installed exclusively in the x3850 X5 (no MAX5).
- 1.5 TB of memory with1 TB in the x3850 X5 server and 0.5 TB installed in the MAX5 memory drawer attached to the x3850 X5.

Data warehouse workloads are host-based environments. No clients were involved to generate the workload. The server ran Microsoft Windows Server 2008 R2 x64 Enterprise Edition.

As to the database layout on the disks, each of the seven SSD PCIe adapters shows up as two drives in Windows, for a total of 14 drives. The SQL database was spread out over all of the 14 drives. The database log was configured on a RAID-10 logical unit number (LUN) built-in six internal SAS drives. The Windows Server 2008 R2 operating system was configured on a 2-disk RAID-1 LUN on internal SAS drives. SQL tempdbqwe was also spread out on the 14 SSD drives.

Analysis and test results

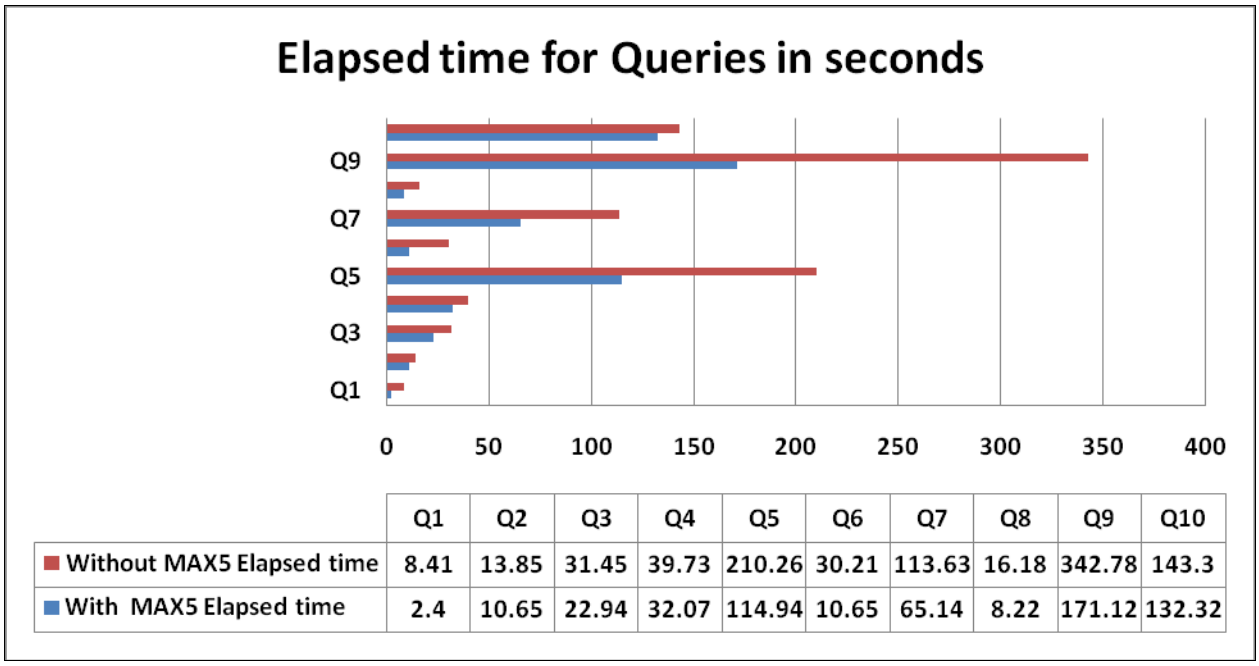
The section provides test results and graphs to show the many benefits of MAX5 and additional memory. When higher amounts of memory are added to the system, thereby enabling the database to be cached in memory, the query response times are greatly improved. The improved query response time can result in reduced I/O and increased CPU utilization. A benefit of using SSD PCIe adapters and large amounts of memory is that the processor utilization is maximized to make the best use of the server.

Sequential query execution

Sample queries were run sequentially to generate the results. The queries are a mix of CPU-intensive and I/O-intensive queries. Certain queries touch just a few rows in the database, but others touch many rows. Certain queries do a full scan through a table, but others are more compute intensive.

The outcome is mixed results, with certain queries running in just a few seconds and others taking several minutes to complete. The same set of queries were run in the same order on both configurations.

Figure 5 shows the comparison of query response times for each of the queries with 1 TB and 1.5 TB of server memory.



	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Without MAX5 Elapsed time	8.41	13.85	31.45	39.73	210.26	30.21	113.63	16.18	342.78	143.3
With MAX5 Elapsed time	2.4	10.65	22.94	32.07	114.94	10.65	65.14	8.22	171.12	132.32

Figure 5 Query elapsed times in seconds

Figure 6 shows the comparison of queries executed per hour during the sequential execution of the queries with 1 TB and 1.5 TB of server memory. The additional memory resulted in a much higher number of queries executed per hour.

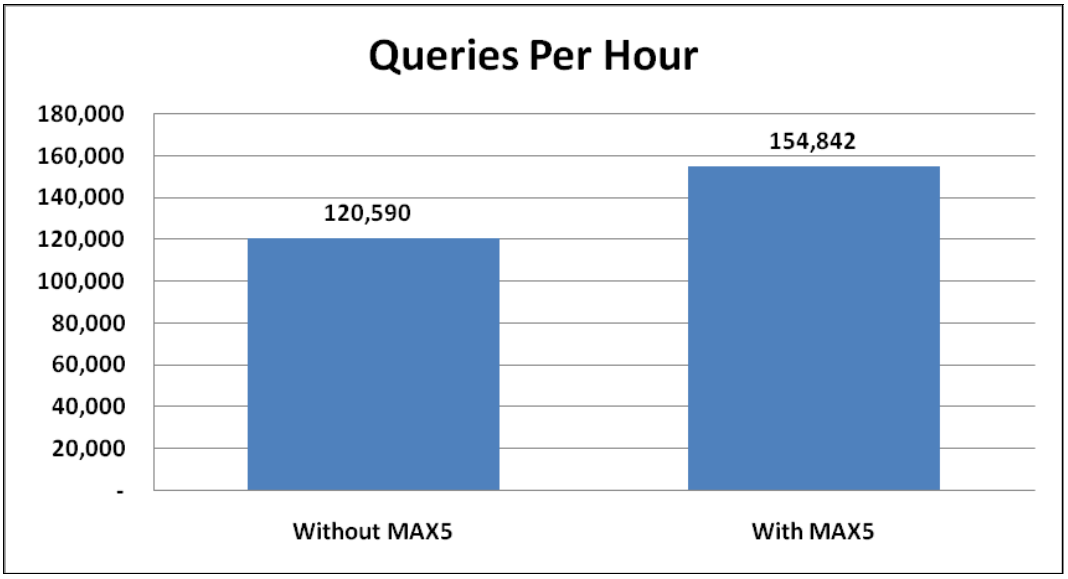


Figure 6 Queries executed per hour

Figure 7 shows the comparison of I/O operations per second during the sequential execution of the queries with 1 TB and 1.5 TB of server memory. With more memory, the database is cached, thereby reducing I/O.

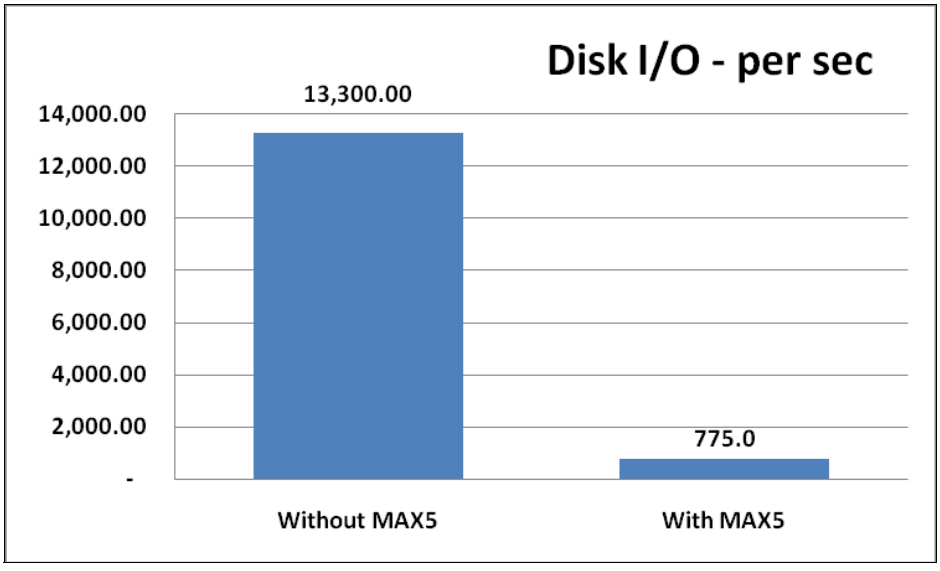


Figure 7 Physical disk I/O per second

Figure 8 shows the comparison of processor utilization percentage during the sequential execution of the queries with 1 TB and 1.5 TB of server memory. Considering that I/O to the PCIe SSD adapters is reduced with MAX5, processor utilization increases resulting in higher query throughput.

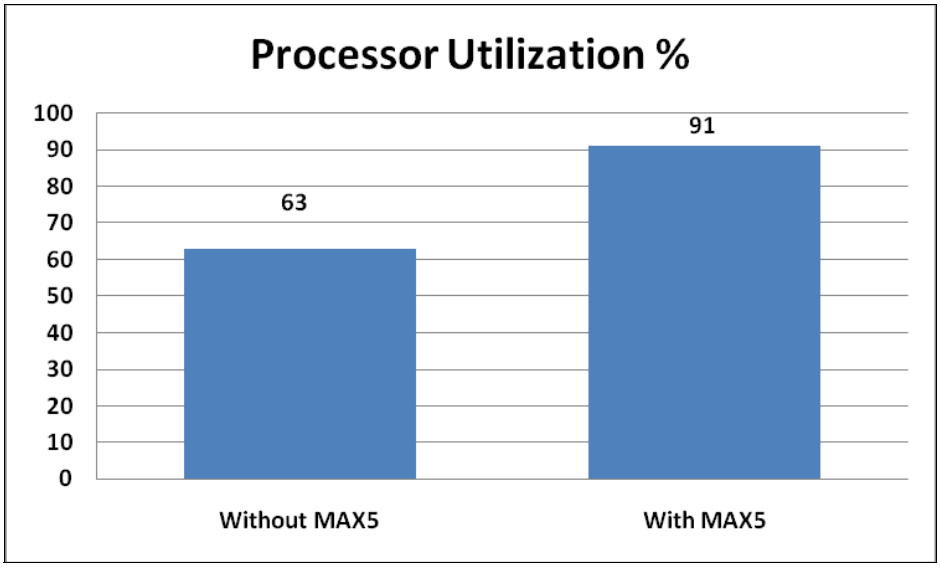


Figure 8 Processor utilization percentages

Parallel query execution

The results in Figure 9 were gathered by issuing the sample queries in parallel.

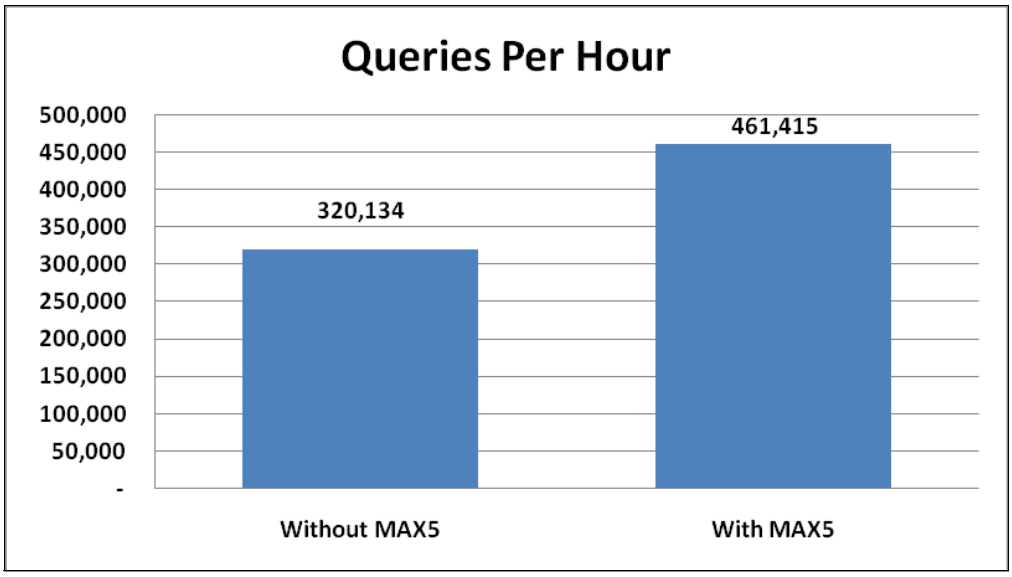
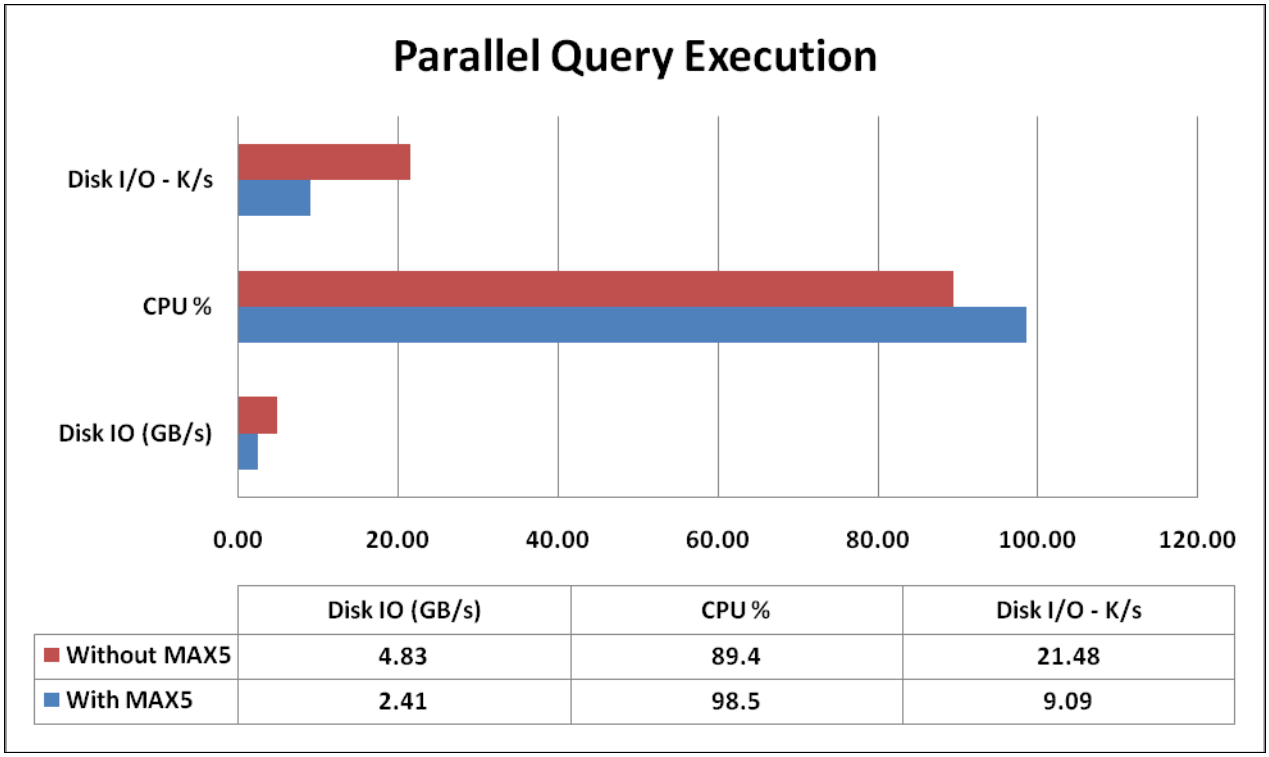


Figure 9 Queries issued per hour during parallel query execution

Figure 9 on page 9 and Figure 10 show data from the tests in which a number of queries were executed in parallel. This data also exhibits the performance benefit of MAX5 and additional memory. Disk I/O decreases, CPU utilization increases, and the number of queries executed per hour increases significantly.



	Disk IO (GB/s)	CPU %	Disk I/O - K/s
Without MAX5	4.83	89.4	21.48
With MAX5	2.41	98.5	9.09

Figure 10 Parallel query execution data

Conclusion

The I/O subsystem is the slowest component in a data warehouse configuration. When you run a query against a database, the server needs to process data. Data resides in the storage, and reads are issued to gather the data set. The IBM High IOPs PCIe adapters help load the data from storage quickly. The SSD PCIe adapters also help with the performance of tempdb activity during the run.

On a system without adequate amounts of system memory, the data that resides in memory must be removed to make room for new data. On a system with a large amount of memory, more of the data set can remain in memory, which enables future references to the data set to be satisfied at low latencies, cutting down the query response times significantly. If achieving faster query response times is your goal and I/O is the bottleneck, then the I/O bandwidth can be increased by either adding more storage or more memory. By adding memory, you can reduce the amount of storage, which in turn offsets the cost of adding that memory by reducing the footprint, space, power, and cooling needs of storage in the data center.

IBM has an effective solution: the 4-socket x3850 X5, which supports up to 1 TB system memory, and up to 1.5 TB of total memory when the MAX5 memory drawer is attached. As presented in this paper, you can gain up to 30 to 40% performance benefit by caching most of the 1.5 TB data warehouse in memory. You can achieve this level performance on an IBM System x3850 X5 with the MAX5 memory drawer attached, and a total of 96 16 GB DIMMs, for a maximum of 1.5 TB of memory in a single four-socket server.

The team who wrote this paper

Vinay Kulkarni is an IBM System x performance engineer working on-site at the Microsoft Redmond campus. He started his career on the AS/400® File system test team, in Rochester, Minnesota. Then he worked on the IBM AS/400 performance benchmark team, publishing many industry leading TPC-C benchmark results. He has been working with Microsoft over the past eight years to optimize performance of System x servers running Microsoft Windows and SQL Server software. Vinay works closely with the Microsoft Windows and SQL Server performance teams to ensure good performance of IBM System x servers running software from Microsoft. He also works with IBM clients to tune the performance of System x and storage environments. He has published leading TPC-H benchmarks recently and works closely with System x marketing team to publish meaningful proof-points based on Microsoft Technologies. Vinay has a Bachelor of Computer Engineering degree from Shivaji University in India. He studied his masters in computer science at the University of Minnesota.

The author would like to acknowledge Tom Scribner of Fusion-io, Inc., for providing the cards to test with, and Michelle Brunk and Jay Holbrook of IBM for supporting the effort.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks
- Follow us on Twitter:

  http://twitter.com/ibmredbooks
- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806
- Explore new Redbooks® publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm
- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document REDP-4751-00 was created or updated on May 5, 2011.

Send us your comments in one of the following ways:
- Use the online **Contact us** review Redbooks form found at:
  ibm.com/redbooks
- Send your comments in an email to:
  redbooks@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks